

HEALTH POLICY

Differences in Physician Income by Gender in a Multiregion Survey

Eric A. Apaydin, PhD^{1,2}, Peggy G. C. Chen, MD, MSc, MHS², and Mark W. Friedberg, MD, MPP³¹Pardee RAND Graduate School, Santa Monica, CA, USA; ²RAND Corporation, Santa Monica, CA, USA; ³RAND Corporation, Boston, MA, USA.

BACKGROUND: Previous studies have documented income differences between male and female physicians. However, the implications of these differences are unclear, since previous studies have lacked detailed data on the quantity and composition of work hours. We sought to identify the sources of these income differences using data from a novel survey of physician work and income.

OBJECTIVE: To compare differences in income between male and female physicians.

DESIGN: We estimated unadjusted income differences between male and female physicians. We then adjusted these differences for total hours worked, composition of work hours, percent of patient care time spent providing procedures, specialty, compensation type, age, years in practice, race, ethnicity, and state and practice random effects.

PARTICIPANTS: We surveyed 656 physicians in 30 practices in six states and received 439 responses (67% response rate): 263 from males and 176 from females.

MAIN MEASURE: Self-reported annual income.

KEY RESULTS: Male physicians had significantly higher annual incomes than female physicians (mean \$297,641 vs. \$206,751; difference \$90,890, 95% CI \$27,769 to \$154,011) and worked significantly more total hours (mean 2470 vs. 2074; difference 396, 95% CI 250 to 542) and more patient care hours (mean 2203 vs. 1845; difference 358, 95% CI 212 to 505) per year. Male physicians were less likely than female physicians to specialize in primary care (49.1 vs. 70.5%), but more likely to perform procedures with (33.1 vs. 15.5%) or without general anesthesia (84.3 vs. 73.1%). After adjustment, male physicians' incomes were \$27,404 (95% CI \$3120 to \$51,688) greater than female physicians' incomes.

CONCLUSIONS: Adjustment for multiple possible confounders, including the number and composition of work hours, can explain approximately 70% of unadjusted income differences between male and female physicians; 30% remains unexplained. Additional study and dedicated efforts might be necessary to identify and address the causes of these unexplained differences.

KEY WORDS: workforce; disparities; reimbursement; health services research; survey research.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11606-018-4462-2>) contains supplementary material, which is available to authorized users.

Received October 19, 2017

Revised March 7, 2018

Accepted April 4, 2018

Published online May 11, 2018

J Gen Intern Med 33(9):1574–81

DOI: 10.1007/s11606-018-4462-2

© Society of General Internal Medicine 2018

INTRODUCTION

Almost 55 years after the passage of the Equal Pay Act, women still earn less than men.¹ Even though approximately half of medical school matriculants are women,^{2, 3} this pay disparity extends to physicians as well.

Prior analyses have noted income differences between men and women in the physician population at large,^{4, 5} among Medicare-participating physicians,⁶ among academic physicians,⁷ and among physicians within the same specialties and geographic areas.³ These income differences have been partially explained by differences in physician age, approximate hours worked, geographic location,⁵ year, and race.⁴ They have also been partially explained using Medicare data on specialty, years of experience, annual number of distinct Medicare beneficiaries served, and annual Medicare service volume.⁶ In addition, a study on academic physicians partially accounted for these differences by age, faculty rank, years since residency, specialty, receipt of National Institutes of Health funding, clinical trial participation, publication count, Medicare reimbursements, graduating medical school rank, and fixed effects for medical schools.⁷

Similar income gaps have been studied for among masters of business administration (MBA) graduates,⁸ lawyers,⁹ and pharmacists.¹⁰ For business professionals with MBAs, the earning gap between men and women progressively widened through 10 years after graduation, although at 10 years, that gap could be explained by pre-MBA characteristics, MBA performance, hours worked, any period of not working, and reasons for choosing one's job. For lawyers, the income gap 15 years after graduation was partially explained in two different law school cohorts by demographics, law school grades, family characteristics, annual work hours, full-time vs. part-time job experience, and job setting.⁹ The gender income gap among pharmacists was large on an annual basis, but was almost fully explained by hours worked.¹⁰ It was theorized that these differences between professions are largely due to non-linear relationships between work hours and income.¹¹ Among MBA holders and lawyers, for example, professionals that work more hours without periods of not working may strengthen client relationships and gain institutional knowledge in ways that are not possible for part-time workers. Over

time, these characteristics enhance productivity, and therefore pay, in ways that cannot be matched through part-time work. On the other hand, pharmacy is a field largely driven by hourly work (nearly 40% of pharmacists earned overtime pay in 2014),¹² and there are fewer opportunities to gain non-linear increases in income. It is not surprising that the gender income gap among pharmacists is almost entirely driven by hours worked. The gender income gap among physicians has not yet been entirely explained by hours worked, and it is still unclear whether the pay structure in the profession more closely resembles that of pharmacy or that of business and law.

Previous studies of the gender income gap in medicine have not controlled for the exact quantity and composition of hours worked. This is not only important because of the possible non-linear relationship between hours and income, but also because not all work hours lead to the same remuneration. For example, time spent performing procedures tends to be more highly compensated than time spent performing cognitive services.¹³

To determine whether controlling for the quantity and composition of work hours could account for observed disparities in incomes, we analyzed responses to a multistate survey of physicians.

METHODS

Data Source

In a 2013 study of physician professional satisfaction, we surveyed 656 practicing physicians in 30 practices within each of six states: Colorado, Massachusetts, North Carolina, Texas, Washington, and Wisconsin.¹⁴ We consulted with each state's medical society and created a list of practices for potential inclusion in the study. Practices were selected to achieve diversity in practice size (10 small with < 10 physicians, 11 medium with 10–49 physicians, 9 large with > 49 physicians), specialty (15 multispecialty, 10 primary care, 5 single subspecialty), and ownership model (19 physician owned or physician partnership, 11 hospital or other corporate ownerships). The practice sample was not nationally representative. Most practices agreed to participate and all but one participating practice completed the study. That practice was replaced before data collection with another practice from the same state. We surveyed all physicians within each practice, receiving 452 responses, of which 439 included a response for physician gender (69% response rate overall; 67% response rate for gender).

The physician survey (see Online Appendix 1 for survey instrument) collected self-reported income from the past year as well as gender, specialty, hours worked per week, weeks worked per year, composition of work hours, percent of patient care time spent providing procedures, compensation type, age, years in practice, race, and ethnicity. This study is a secondary analysis of the survey data; the survey was originally intended to measure physician professional satisfaction

and related determinants. Analysis of these survey data was approved by the RAND Human Subjects Protection Committee.

Measures

Our outcome was annual income in dollars. Demographic covariates included the following: gender (male/female), specialty group (primary care: family practice, general practice, internal medicine, and pediatrics; obstetrics/gynecology; medical specialties: cardiology, dermatology, gastroenterology, neurology, oncology, and pulmonology; surgical specialties: general surgery, otolaryngology, ophthalmology, orthopedic surgery, and urology; other: emergency medicine, psychiatry and other), race (white [European, Middle Eastern, other], black or African American, American Indian or Alaskan, Native Hawaiian or Pacific Islander, Asian, or other), ethnicity (Hispanic or Latino, or Not Hispanic or Latino), age (grouped as 18–40 years old, 41–50 years old, 51–60 years old, 61–70 years old, or 71 years old or more), and years in practice (5 years or less, 6–10 years, 11–20 years, 21–30 years, or 31 years or more). Work covariates included the following: hours worked per year (in hours), work hour composition (percentage of time spent on patient care, teaching, research, administration, and other), percent of patient care hours spent performing procedures with and without general anesthesia (categorized as 0%, 1–24%, or $\geq 25\%$), and compensation type (fixed salary, salary adjusted for performance, shift, hourly or other time-based method, or share of practice billing or workload). Practice covariates included the following: practice size (≤ 9 physicians, 10–49 physicians, or ≥ 50 physicians), practice specialty (primary care, single subspecialty, or multispecialty), and practice ownership model (physician-owned or partnership, or hospital or corporate owner). We also included state and practice random effects.

Statistical Analyses

We estimated several multilevel mixed-effects generalized linear models of annual income as a function of different covariates. Model 1 estimated income as a function of physician gender alone; Model 2 additionally adjusted for state and practice random effects and practice covariates; Model 3 additionally adjusted for hours worked per year; Model 4 additionally adjusted for specialty. Model 5 additionally adjusted for work hour composition, and Model 6 additionally adjusted for the percent of patient care hours spent performing procedures with and without general anesthesia, and compensation type. Model 7 additionally adjusted for age, years in practice, race, and ethnicity. We report the difference between male and female incomes in every model to show how the difference was “explained” by additional covariates and how much of the residual difference remained unexplained.

We excluded respondents whose reported work hours exceeded the 95th percentile (3550 h/year) because such extreme values might have been misreported or might be

considered “overtime” hours that deserve higher effective hourly wages. In all models, we winsorized income at the 95th percentile (\$600,000) to limit the influence of extreme outliers. Winsorization of these outliers can help compare “like to like,” as respondents earning very high incomes may have fundamentally different work or business arrangements compared to all other respondents.¹⁵ Non-response can bias estimates of parameters and relationships even among surveys with high response rates.^{16, 17} Therefore, we constructed non-response weights using data on gender, age, AMA membership, state, individual specialty, practice size, practice specialty, and practice ownership model and used the weights in all models to account for differences between survey responders and non-responders. Multiple imputation by chained equations is the preferred method for addressing missing data in observational studies.^{18, 19} Here, we used multiple imputation for the following covariates: gender; age; years in practice; race; ethnicity; hours worked per year; percent of hours worked in teaching, research, administration, and other; percent of patient care hours spent performing procedures with and without general anesthesia; and compensation type. State effects were assigned to the state level of the model. Practice random effects and a vector of practice characteristics (practice specialty, size, and ownership model) were assigned to the practice level of the model. All other covariates were assigned to the individual level. Standard errors were clustered within states and practices.

RESULTS

Sample Characteristics

The survey sample contained 439 physicians with gender data (263 men and 176 women, Table 1). The winsorized mean income in the sample was \$259,838 overall (95% CI \$246,268 to \$273,408) and \$255,991 (95% CI \$242,341 to \$269,640) among those working less than the 95th percentile in annual hours, with men earning significantly more than women in both cases (overall: mean \$296,773 vs. \$201,546; difference \$95,228, 95% CI \$69,062 to \$134,355; excluding respondents with annual work hours above the 95th percentile: mean \$291,279 vs. \$201,701; difference \$89,578, 95% CI \$63,186 to \$115,971). Additional estimates of income overall and between the genders are presented in Table 1.

Men reported working significantly more total hours per year (mean 2470 vs. 2074; difference 396, 95% CI 250 to 542) and having more patient care hours per year (mean 2203 vs. 1845; difference 358, 95% CI 212 to 505) than women. Women worked significantly more hours in other uncategorized activities (mean 33 vs. 12; difference 21, 95% CI 1 to 40) than men. Hours worked in teaching, research, and administration did not differ significantly between genders. A larger percentage of men than women provided procedures, both with general anesthesia (33.1 vs. 15.5%) and without general anesthesia (84.3 vs. 73.1%).

Male physicians were less likely than female physicians to be in primary care specialties (49.1 vs. 70.5%). Mean winsorized incomes among those working less than the 95th percentile in annual hours were lower in primary care (\$196,787) than in obstetrics/gynecology (\$279,417), medical subspecialties (\$411,233), surgical subspecialties (\$427,649), and other subspecialties (\$280,172). Winsorized incomes per specialty were similar for physicians working any number of hours. Also, a larger percentage of male physicians in the sample were white (86.8 vs. 73.9%), older than 40 (85.8 vs. 66.7%), had more than 10 years in practice (74.1 vs. 56.3%), and had non-fixed salaries (86.9 vs. 72.1%). More men worked in single subspecialty practices (20.2 vs. 6.3%) and in practices that were physician-owned or a partnership (52.1 vs. 38.6%). The number of male and female physicians surveyed also significantly differed between states in the sample. There was no significant difference in Hispanic ethnicity or practice size between the groups. Demographically, our sample was of similar age, but more female, more white, less Hispanic, and more specialized in primary care compared to several national samples (Online Appendix 2).^{20, 21}

Multivariable Analyses

Table 2 describes the difference between mean male and female incomes in each of our models and the covariates in each model. All models include winsorized income at the 95th percentile, a sample restricted to the bottom 95% of hours (< 3550 h per year), non-response weights, and imputed non-outcome covariates. In the unadjusted model, model 1, male physicians reported annual incomes that were \$90,890 (95% CI \$27,769 to \$154,011) higher than female physicians. This difference decreased to \$51,308 (95% CI \$9636 to \$92,981) when adjusting for state, practice, and practice characteristics random effects (model 2). The difference decreased further to \$39,521 (95% CI \$6646 to \$72,396) when additionally adjusting for hours worked (model 3). When additionally adjusting for specialty group, this difference fell further, to \$26,535 (95% CI \$526 to \$52,543) (Model 4). In model 5, we further adjusted for the percentage of work hours in each category, with modest marginal reduction in mean adjusted male-female income difference \$24,726 (95% CI \$-1174 to \$50,625). In model 6, men earned \$26,679 (95% CI \$-1048 to \$54,405) more than women after additionally adjusting for percent of patient care time spent providing procedures under and not under general anesthesia, and compensation type.

In model 7, our full model, male physicians earned \$27,404 (95% CI \$3120 to \$51,688) more than female physicians after additionally adjusting for age, years in practice, race, and ethnicity. This indicates that adjusting for the characteristics in the full model explained 69.8% of the unadjusted winsorized income difference between men and women in our sample. Additional model specifications are detailed in Online Appendices 3 and 4.

Table 1 Characteristics of the Sample

Characteristic	Total (n = 439)	Male (n = 263)	Female (n = 176)	P value	n
Income/wage (mean, 95% CI), \$					
Income				<0.001*	379
Winsorized income	273,490 (254,087 to 292,892)	311,316 (287,404 to 335,228)	213,790 (183,027 to 244,554)	<0.001*	379
Winsorized income; annual hours <95th percentile	259,838 (246,268 to 273,408)	296,773 (278,599 to 314,947)	201,546 (185,313 to 217,778)	<0.001*	363
Winsorized income; annual hours <95th percentile; imputed gender	255,991 (242,341 to 269,640)	291,279 (272,843 to 309,714)	201,701 (185,174 to 218,227)	<0.001*	369
Winsorized income; annual hours <95th percentile; imputed gender; non-response weights; state and practice clustered standard errors	257,096 (243,499 to 270,694)	292,057 (273,662 to 310,452)	203,084 (186,280 to 219,888)	0.017*	365
Hours (mean, 95% CI)	262,386 (202,766 to 322,006)	297,641 (236,202 to 359,081)	206,751 (168,377 to 245,125)		
Total hours worked per year	2314 (2241 to 2388)	2470 (2380 to 2559)	2074 (1955 to 2192)	<0.001*	426
Patient care hours per year	2063 (1989 to 2136)	2203 (2114 to 2292)	1845 (1724 to 1966)	<0.001*	426
Hours, cont'd (mean, 95% CI)					
Teaching hours per year	46 (36 to 57)	49 (36 to 63)	42 (25 to 59)	0.515	426
Research hours per year	15 (8 to 22)	15 (6 to 24)	17 (5 to 28)	0.800	426
Administration hours per year	190 (157 to 223)	202 (163 to 242)	170 (111 to 229)	0.357	426
Other hours per year	21 (11 to 30)	12 (3 to 22)	33 (14 to 53)	0.036*	426
Specialty or specialty group (n, column %)					
Specialty: cardiology	14 (3.2%)	11 (4.2%)	3 (1.7%)	0.148	14
Specialty: dermatology	3 (0.7%)	2 (0.8%)	1 (0.6%)	0.811	3
Specialty: emergency medicine	1 (0.2%)	1 (0.4%)	0 (0%)	0.413	1
Specialty: family practice	87 (19.8%)	44 (16.7%)	43 (24.4%)	0.047*	87
Specialty: general practice	8 (1.8%)	2 (0.8%)	6 (3.4%)	0.042*	8
Specialty: general pediatrics	49 (11.1%)	18 (6.8%)	31 (17.6%)	<0.001*	49
Specialty: gastroenterology	6 (1.4%)	4 (1.5%)	2 (1.1%)	0.734	6
Specialty: general surgery	19 (4.3%)	14 (5.3%)	5 (2.8%)	0.210	19
Specialty: internal medicine	109 (24.8%)	65 (24.7%)	44 (25.0%)	0.946	109
Specialty: neurology	3 (0.7%)	3 (1.1%)	0 (0%)	0.155	3
Specialty: oncology	7 (1.6%)	6 (2.3%)	1 (0.6%)	0.160	7
Specialty: obstetrics/gynecology	27 (6.2%)	14 (5.3%)	13 (7.4%)	0.378	27
Specialty: ophthalmology	3 (0.7%)	2 (0.8%)	1 (0.6%)	0.811	3
Specialty: orthopedic surgery	21 (4.8%)	20 (7.6%)	1 (0.6%)	0.001*	21
Specialty: otolaryngology	6 (1.4%)	6 (2.3%)	0 (0%)	0.044*	6
Specialty: psychiatry	3 (0.7%)	2 (0.8%)	1 (0.6%)	0.811	3
Specialty: pulmonology	5 (1.1%)	5 (1.9%)	0 (0%)	0.066	5
Specialty: urology	8 (1.8%)	8 (3.0%)	0 (0%)	0.020*	8
Specialty: other	57 (12.9%)	32 (12.2%)	25 (14.2%)	0.534	57
Specialty group: primary care (family practice, general practice, internal medicine, and pediatrics)	253 (57.6%)	129 (49.1%)	124 (70.5%)	<0.001*	253
Specialty group: obstetrics/gynecology	27 (6.2%)	14 (5.3%)	13 (7.4%)	0.378	27
Specialty group: medical specialties (cardiology, dermatology, gastroenterology, neurology, oncology, and pulmonology)	38 (8.7%)	31 (11.8%)	7 (4.0%)	0.004*	38
Specialty group: surgical specialties (general surgery, otolaryngology, ophthalmology, orthopedic surgery, and urology)	57 (13.0%)	50 (19.0%)	7 (4.0%)	<0.001*	57
Specialty group: other (emergency medicine, psychiatry and other)	61 (13.9%)	35 (13.3%)	26 (14.8%)	0.664	61
Patient care activities (n, column %)				0.001*	434
Percent of patient care hours: providing procedures to patients who are under general anesthesia					
Never	321 (74.0%)	174 (66.9%)	147 (84.5%)		
1–9%	36 (8.2%)	23 (8.9%)	13 (7.5%)		
10–24%	29 (6.7%)	24 (9.2%)	5 (2.9%)		
25–49%	20 (4.6%)	15 (5.8%)	5 (2.9%)		

(continued on next page)

Table 1. (continued)

Characteristic	Total (n = 439)	Male (n = 263)	Female (n = 176)	P value	n
50–74%	23 (5.3%)	19 (7.3%)	4 (2.3%)		
> 75%	5 (1.2%)	5 (1.9%)	0 (0%)	0.017*	416
Percent of patient care hours: providing procedures to patients who are not under general anesthesia					
Never	84 (20.2%)	39 (15.7%)	45 (26.9%)		
1–9%	200 (48.1%)	117 (47.0%)	83 (49.7%)		
10–24%	76 (18.3%)	52 (20.9%)	24 (14.4%)		
25–49%	27 (6.5%)	19 (7.6%)	8 (4.8%)		
50–74%	17 (4.1%)	14 (5.6%)	3 (1.8%)		
> 75%	12 (2.9%)	8 (3.2%)	4 (2.4%)		
Compensation type (n, column %)					
Fixed salary	78 (18.9%)	33 (13.1%)	45 (28.0%)	< 0.001*	413
Salary adjusted for performance	303 (73.4%)	190 (75.4%)	113 (70.2%)		
Shift, hourly or other time-based method	11 (2.7%)	11 (4.4%)	0 (0%)		
Share of practice billings or workload	21 (5.1%)	18 (7.1%)	3 (1.9%)		
Age (n, column %)					
18–30 years old	0 (0.0%)	0 (0.0%)	0 (0.0%)		
31–40 years old	94 (21.8%)	37 (14.2%)	57 (33.3%)		
41–50 years old	147 (34.0%)	88 (33.7%)	59 (34.5%)		
51–60 years old	129 (29.9%)	85 (32.6%)	44 (25.7%)		
61–70 years old	54 (12.5%)	44 (16.9%)	10 (5.9%)		
71 years old or more	8 (1.9%)	7 (2.7%)	1 (0.6%)		
Years in practice (n, column %)					
5 years or less	69 (15.7%)	26 (9.9%)	43 (24.4%)	< 0.001*	439
6–10 years	76 (17.3%)	42 (16.0%)	34 (19.3%)		
11–20 years	129 (29.4%)	76 (28.9%)	53 (30.1%)		
21–30 years	111 (25.3%)	73 (27.8%)	38 (21.6%)		
31 years or more	54 (12.3%)	46 (17.5%)	8 (4.6%)		
Race (n, column %)					
White (European, Middle Eastern, other)	345 (81.8%)	223 (86.8%)	122 (73.9%)	0.018*	422
Black or African American	9 (2.1%)	3 (1.2%)	6 (3.6%)		
American Indian or Alaskan	3 (0.7%)	1 (0.4%)	2 (1.2%)		
Native Hawaiian or Pacific Islander	2 (0.5%)	0 (0%)	2 (1.2%)		
Asian	59 (14.0%)	28 (10.9%)	31 (18.8%)		
Other	4 (1.0%)	2 (0.8%)	2 (1.2%)		
Ethnicity (n, column %)					
Not Hispanic or Latino	409 (96.5%)	247 (96.5%)	162 (96.4%)	0.976	427
Hispanic or Latino	15 (3.5%)	9 (3.5%)	6 (3.6%)	< 0.001*	438
State (n, column %)					
CO	53 (12.1%)	38 (14.5%)	15 (8.5%)		
MA	94 (21.5%)	31 (11.8%)	63 (35.8%)		
NC	76 (17.4%)	56 (21.4%)	20 (11.7%)		
TX	59 (13.5%)	36 (13.5%)	23 (13.1%)		
WA	63 (14.4%)	41 (15.7%)	22 (12.5%)		
WI	93 (21.2%)	60 (22.9%)	33 (18.8%)		
Practice specialty (n, column %)					
Primary care	38 (8.7%)	18 (6.8%)	20 (11.4%)	< 0.001*	439
Single subspecialty	64 (14.6%)	54 (20.2%)	11 (6.3%)		
Multispecialty	337 (76.8%)	192 (73.0%)	145 (82.4%)		

(continued on next page)

Table 1. (continued)

Characteristic	Total (n = 439)	Male (n = 263)	Female (n = 176)	P value	n
Ownership model (n, column %)					
Physician-owned or partnership	205 (46.7%)	137 (52.1%)	68 (38.6%)	0.006*	439
Hospital or corporate owned	234 (53.5%)	126 (47.9%)	108 (61.4%)		
Practice size (n, column %)					
≤ 9 physicians	28 (6.4%)	17 (6.5%)	11 (6.3%)	0.230	439
10–49 physicians	188 (42.8%)	121 (46.0%)	67 (38.1%)		
≥ 50 physicians	223 (50.8%)	125 (47.3%)	98 (55.7%)		

* $P < 0.05$; P values reflect two-sided t tests or chi-squared tests for continuous and binary or categorical variables, respectively

DISCUSSION

We analyzed gender differences in physician income using a survey administered in 30 practices across six states. Our adjusted models could account for approximately 70% of the unadjusted difference in incomes between male and female physicians.

The residual, unexplained income gap between male and female physicians might represent disparities in income due to differences in work composition that were not captured by our survey (e.g., performing more profitable procedures), or it may represent discrimination in pay. Building more complete models of the physician gender income gap and determining the composition of unexplained residuals might be important to understanding and closing gaps in physician pay. Of note, we were able to reduce this residual by more than in a prior analysis of a national sample of physicians and by an amount similar to that in a prior analysis of academic physicians (Online Appendix 5).^{4,7} The national study of physicians only employed limited controls for demographics (age, race, geographic location, year) and productivity (approximate hours worked) compared to our analysis.⁴ It is therefore expected that our models explain more of the income gap, even with a smaller sample size. Similar to our analysis, the prior study of academic physicians partially explained the residual income gap using demographic factors (age, faculty rank, years since residency), specialty, measures of productivity (receipt of National Institutes of Health funding, clinical trial participation, publication count, Medicare reimbursements), and measures of quality (graduating medical school rank, fixed effects for medical schools).⁷ These factors mirror the covariates in our model, with the exception of measures of physician quality. Still, the persistence of similar unexplained residuals in models of separate populations of physicians with separate covariates suggests that other drivers of gender differences in income are present. Future work should employ longitudinal analyses to examine the change in the gender income over physicians' careers, especially with respect to periods of part-time work or non-work, as with prior analyses among business professionals and lawyers.^{8,9} These analyses should also examine hourly earnings in a national sample, similar to a prior study of pharmacists.¹⁰ Understanding the differences in income between male and female physicians over time is likely a key to understanding the drivers of the gap. The unexplained residual in the gap could be driven by non-linear increases in productivity, as described in studies of business professional or lawyers, or by other factors, including overt discrimination, types of procedures performed, or patient volume. Future research should consider all of these factors in studying the gender income gap among physicians.

Our study has limitations. First, our analysis was based on responses to a survey that was of modest sample size and which was not designed to be nationally representative. These survey characteristics limit the generalizability of our findings. However, our findings are similar to those of nationally representative samples,^{4,7} and our sample of survey responders is similar to non-responders on several demographic characteristics.

Table 2 Summary of Multivariable Analyses of Characteristics Associated with Physician Income

Model	Male/female difference in income, coefficient (95% CI), \$
Model 1: no adjustment	90,890 (27,769 to 154,011)
Model 2: state random effects + practice random effects + practice characteristics random effects	51,308 (9636 to 92,981)
Model 3: hours worked per year + state random effects + practice random effects + practice characteristics random effects	39,521 (6646 to 72,396)
Model 4: hours worked per year + specialty group + state random effects + practice random effects + practice characteristics random effects	26,535 (526 to 52,543)
Model 5: hours worked per year + percent of hours worked per category + specialty group + state random effects + practice random effects + practice characteristics random effects	24,726 (-1174 to 50,625)
Model 6: hours worked per year + percent of hours worked per category + specialty group + percent of patient care time spent performing procedures under and not under anesthesia + compensation type + state random effects + practice random effects + practice characteristics random effects	26,679 (-1048 to 54,405)
Model 7: hours worked per year + percent of hours worked per category + specialty group + percent of patient care time spent performing procedures under and not under anesthesia + compensation type + age + years in practice + race + ethnicity + state random effects + practice random effects + practice characteristics random effects	27,404 (3120 to 51,688)

All models include Winsorized income at the 95th percentile and are restricted to observations with less than the 95th percentile of annual hours worked (< 3550 h per year) to avoid the influence of outliers. They also include non-response weights and imputed independent variables to avoid the influence of non-response and missing data. Models 2 through 7 additionally include standard errors clustered by state and practice

Second, our dataset contained missing data for self-reported income and for several explanatory covariates. To correct for this, we used multiple imputation by chained equations to impute missing data for non-outcome covariates. This imputation ensures that all models have the same sample size and are comparable. Third, we relied on self-reported incomes and hours worked per year; we were unable to validate these figures by comparison to another data source. Fourth, we lacked data on patient volume or type of procedures performed.

Men reported significantly higher incomes than women in a sample of 439 physicians in 30 practices across six states. After adjusting for a variety of relevant factors, including hours worked and time spent performing procedures, approximately 30% of this residual difference continued to be unexplained. Future work is necessary to determine the cause of this persistent disparity.

Acknowledgements: We would like to acknowledge Dr. Bonnie Ghosh-Dastidar, Mr. Aaron Kofner, and Mr. Q Burkhardt for their advice on multiple imputation and non-response analysis during the revision process of this manuscript.

Corresponding Author: Eric A. Apaydin, PhD; RAND Corporation, Santa Monica, CA, USA (e-mail: eapaydin@rand.org).

Funding Dr. Apaydin received partial support for this research from the Pardee RAND Graduate School Anne and James Rothenberg Dissertation Award. The content herein is solely the responsibility of the authors and does not necessarily reflect the views of the Pardee RAND Graduate School or the RAND Corporation.

Compliance with Ethical Standards:

Conflict of Interest: The authors declare that they do not have a conflict of interest.

REFERENCES

1. Proctor BD, Semega JL, Kollar MA. Income and poverty in the United States: 2015. U.S. census bureau. 2016. Available from: <http://www.census.gov/library/publications/2016/demo/p60-256.html>.
2. Association of American Medical Colleges. 2016 applicant and enrollment data charts. 2016. Available from: http://aamc-black.global.ssl.fastly.net/production/media/filer_public/78/3e/783e28cc-4bbf-4fab-a23a-21f990cd2683/20161006_applicant_and_matriculant_data_tables.pdf.
3. Whaley C. First annual physician compensation report. San Francisco: Doximity. 2017. Available from: http://s3.amazonaws.com/s3.doximity.com/careers/2017_physician_compensation_report.pdf.
4. Ly DP, Seabury SA, Jena AB. Differences in incomes of physicians in the United States by race and sex: observational study. *BMJ* 2016;353:i2923.
5. Seabury SA, Chandra A, Jena AB. Trends in the earnings of male and female health care professionals in the United States, 1987 to 2010. *JAMA Intern Med* 2013;173(18):1748–50.
6. Desai T, Ali S, Fang X, Thompson W, Jawa P, Vachharajani T. Equal work for unequal pay: the gender reimbursement gap for healthcare providers in the United States. *Postgrad Med J* 2016;92(1092):571–5.
7. Jena AB, Olenski AR, Blumenthal DM. Sex differences in physician salary in US public medical schools. *JAMA Intern Med* 2016;176(9):1294–304.
8. Bertrand M, Goldin C, Katz LF. Dynamics of the gender gap for young professionals in the financial and corporate sectors. *Am Econ J: Appl Econ* 2010;2(3):228–55.
9. Noonan MC, Corcoran ME, Courant PN. Pay differences among the highly trained: cohort differences in the sex gap in lawyers' earnings. *Social Forces* 2005;84(2):853–72.
10. Goldin C, Katz L. The most egalitarian of all professions: pharmacy and the evolution of a family-friendly occupation. NBER Working Paper Series. 2012. Available from: <http://www.nber.org/papers/w18410>.
11. Goldin C. A Grand gender convergence: its last chapter. *Am Econ Rev* 2014;104(4):1091–119.
12. Midwest Pharmacy Workforce Research Consortium. 2014 national pharmacist workforce survey. Minneapolis, MN. 2015. Available from: <http://www.aacp.org/sites/default/files/executivesummaryfromthenationalpharmacistworkforcestudy2014.pdf>.
13. Sinsky CA, Dugdale DC. Medicare payment for cognitive vs procedural care: minding the gap. *JAMA Intern Med*. 2013;173(18):1733–7.
14. Friedberg MW, Chen PG, Van Busum KR, Aunon F, Pham C, Caloyeras J, et al. Factors affecting physician professional satisfaction and their implications for patient care, health systems, and health policy. Santa

- Monica, CA: RAND Corporation. 2013. Available from: http://www.rand.org/pubs/research_reports/RR439.
15. **Spiegelhalter D, Sherlaw-Johnson C, Bardsley M, Blunt I, Wood C, Grigg O.** Statistical methods for healthcare regulation: rating, screening and surveillance. *J R Stat Soc A (Stat Soc)* 2012;175(1):1–47.
 16. **McLeod CC, Klabunde CN, Willis GB, Stark D.** Health care provider surveys in the United States, 2000–2010: a review. *Eval Health Prof* 2013;36(1):106–26.
 17. **Johnson TP, Wislar JS.** Response rates and nonresponse errors in surveys. *JAMA* 2012;307(17):1805–6.
 18. **Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al.** The prevention and treatment of missing data in clinical trials. *New England Journal of Medicine*. 2012;367(14):1355–60.
 19. **Ware JH, Harrington D, Hunter DJ, D'Agostino RB.** Missing data. *New England Journal of Medicine*. 2012 367(14):1353–4.
 20. **Association of American Medical Colleges.** 2016 physician specialty data report. Washington, DC. 2016. Available from: <http://www.aamc.org/data/workforce/reports/457712/2016-specialty-databook.html>.
 21. **Association of American Medical Colleges.** Diversity in the physician workforce: facts & figures 2010. Washington, DC. 2010. Available from: <http://www.aamc.org/download/432976/data/factsandfigures2010.pdf>.